# Combining Analytical Model and Experimental Test Data for Optimal Determination of Failure Tolerance Limits

**Stanislav Uryasev**
American Optimal Decisions, Inc.
and
Department of Industrial & Systems Engineering
University of Florida
Gainesville, Florida, USA
*uryasev@ufl.edu*

**A. Alexandre Trindade**
Department of Statistics
University of Florida
Gainesville, Florida, USA
*trindade@stat.ufl.edu*

## Abstract

We propose a coherent methodology for integrating various sources of variability on properties of materials, in order to accurately predict percentiles of their failure load distribution. The approach involves the linear combination of appropriate factors into a regression model. "CVaR Deviation", a concept borrowed from financial risk management, is used as the optimality criterion in model-fitting. The approach yields identical results to Quantile Regression, but is more general in scope. We consider confidence interval construction methods for the percentiles, and adopt the most promising of these to compute statistical tolerance limits for the failure load. The approach pools together information from physical tests and analytical model runs, resulting in accurate inferences even in the presence of relatively small datasets. Simulations suggest the approach is quite robust, and allows for a substantial reduction in the number costly physical tests that have to be performed, with little sacrifice in precision.

## 1 The Problem

A basic problem in engineering design is how to manage uncertainty and risk in the properties of materials. Albeit costly, physical testing of the materials has traditionally been the primary method of quantifying the uncertainty. Recent advances in analytical physics-based engineering models and associated hardware/software technology, have made this an increasingly important contributor. Although potentially less precise, the lower cost of the **analytical method** makes it an attractive contender to the **testing method**. A sound risk management strategy in this context might argueably be the effective integration of the information obtained from each method.

The purpose of this paper is to propose a coherent methodology for the integration of heterogenous sources of information, by combining data into a regression factor model. Such an approach will necessarily have to allow for treating of the test data as both covariate and response. Desirable inferences from the resulting procedure will take the form of point and interval estimates of percentiles of material properties. Special cases include failure load tolerance limits such as A-Basis and B-Basis values[1].

## 2 The Method

### 2.1 Model Setup

The basic idea is to use each of the test data values as the response in turn; the remaining test data and all the analytical model data serving as explanatory variables. Let $Y_{ij}$ denote the $j$th

[1] An A-Basis value for the random variable $X$ is a 95% lower confidence bound for the 1st percentile of $X$ (in repeated sampling the calculated A-Basis value would fall below the true 1st percentile 95% of the time). A B-Basis value is defined similarly for the 10th percentile of $X$.

test data point corresponding to the $i$th batch[2], $i = 1, \ldots, I$, and $j = 1, \ldots, n_i$. We condense the data from each source, test and model, into a pair of summary statistics such as the mean and standard deviation. Other data reduction measures of location (e.g. median) and dispersion (e.g. lower semi-deviation) could also be used for datasets with asymmetrical distributions. Letting $(m_i, s_i)$ and $(\mu_i, \sigma_i)$ denote the sample mean and standard deviation for the test and model data respectively in batch $i$, we fit the regression model:

$$Y_{ij} = c_0 + c_1\mu_i + c_2\sigma_i + c_3 m_i + c_4 s_i + \varepsilon_{ij}, \tag{1}$$

where $c_0, \ldots, c_4$ are unknown regression coefficients to be estimated from the data, and $\varepsilon_{ij}$ is the residual error corresponding to the $j$th test data point in the $i$th batch.

### 2.2 Model Fitting

Let $\varepsilon$ denote the residual error distribution, with cdf $F^{-1}$, in the above regression model. The deviation CVaR (Rockafellar, Uryasev, and Zabarankin, 2002) at probability level $\tau$, is defined as

$$\mathrm{CVaR}^{\Delta}_{\tau}(\varepsilon) \;\equiv\; \mathbf{E}\Big[\, \varepsilon - \mathbf{E}\varepsilon \mid \varepsilon \geq F^{-1}(\tau)\Big] \;, \qquad \tau > 0.5 \;\; .$$

This is simply the average distance between the mean value and the values in the $\tau$ tail of the distribution, and as such, is a one-sided measure of the width of such distribution. For the lower tail, deviation CVaR is defined analogously as

$$\mathrm{CVaR}^{\Delta}_{1-\tau}(-\varepsilon) \;\equiv\; \mathbf{E}\Big[\, \mathbf{E}\varepsilon - \varepsilon \mid \varepsilon \leq F^{-1}(\tau)\Big] \;, \qquad \tau \leq 0.5 \;\; .$$

Deviation CVaR is now the average distance between the mean value and the values in the lower $\tau$ tail of the distribution.

The fitting of the parameters in regression model (1) then entails minimizing the following optimality criterion over all $(c_1, \ldots, c_4) \in \mathbf{R}^4$:

$$\mathcal{P}_{\tau}(\mathbf{c}) = \begin{cases} \mathrm{CVaR}^{\Delta}_{\tau}(\varepsilon)\,, & \text{if } \tau \geq 0.5\,, \\ \mathrm{CVaR}^{\Delta}_{1-\tau}(-\varepsilon)\,, & \text{if } \tau < 0.5\,. \end{cases} \tag{2}$$

This leads to the following estimated equation for the $\tau$th quantile of the failure load, as a function of the generic factors $\{\mu, \sigma, m, s\}$:

$$\hat{Q}_Y(\tau) = \hat{c}_0(\tau) + \hat{c}_1(\tau)\mu + \hat{c}_2(\tau)\sigma + \hat{c}_3(\tau)m + \hat{c}_4(\tau)s. \tag{3}$$

Actually, the coefficient $\hat{c}_0(\tau) = \mathrm{VaR}_{\tau}(\hat{\varepsilon}(\tau))$ is not determined by the minimization of $\mathcal{P}_{\tau}(\mathbf{c})$, but can for example be chosen so that $\hat{Q}_Y(\tau)$ is an unbiased quantile estimate, where $\hat{\varepsilon}(\tau)$ is the residual in an optimal point. This leads to the same estimates as in quantile regression, introduced by Koenker and Bassett (1982). The optimization of $\mathcal{P}_{\tau}(\mathbf{c})$ can be efficiently performed via linear programming (Uryasev and Rockafellar 2000, 2002). The attained minimum value of $\mathcal{P}_{\tau}(\mathbf{c})$ is a natural candidate for a measure of goodness-of-fit.

An important advantage of this **CVaR regression** method, is that only mild distributional assumptions are made, making the approach essentially nonparametric. A-basis and B-basis values for a generic covariate vector $[1, \mu_i, \sigma_i, m_i, s_i]$, are obtained as a 95% lower confidence bound for the respective quantile, $\tau = 0.01$ or $\tau = 0.10$. These confidence bounds can be obtained via the method outlined in Koenker (1994), using for example the Hall and Sheather bandwidth.

[2]The term "batch" will denote any extraneous source of variability affecting the material; materials from different batches assumed *a priori* to have different properties.

## 3 Simulation Results

The performance of the CVaR regression method was assessed in a controlled environment consisting of failure data simulated from Weibull distributions, with various choices for the shape and scale parameters. We assess goodness-of-fit by measuring Mean Absolute Deviation (MAD) between true and estimated 10th quantiles across batches, and Mean CVaR Deviation (MCD) for CVaR regression fits. The CVaR deviation is that given by equation (2). A particular value of MAD or MCD is based on 100 draws from the same Weibull distribution. The choice of shape ($\alpha$) and scale ($\beta$) parameters, constituting a batch, is made randomly and uniformly over the rectangle, $10 < \alpha < 80$ and $40 < \beta < 120$. Each draw is of sample size $n + 100$. The first $n$ points are then used as test data, and the remaining 100 as model data. A new set of shape and scale parameters is then randomly selected to produce data on the next batch, etc. Robustness of the methodology was also assessed by perturbing the scale and shape parameters of the model data with respect to the test data parameters (the true distribution).

Figure 1 shows the resulting deviations plotted as functions of number of batches ($m$) used in the CVaR regression fit, as well as number of test points ($n$) within in each batch. The main findings can be summarized as follows:

(i) MAD decreases as the number of batches and/or test points per batch increases.

(ii) MCD increases as the number of batches and/or test points per batch increases.

(iii) The accuracy of CVaR regression is relatively insensitive to the number of batches present; but fairly sensitive to the number of test points per batch.

(iv) There are diminishing benefits in using more than 10 batches, or more than 10 test points per batch, in any one experiment.

(vi) MAD and MCD are virtually unaffected by shape biases; the methodology therefore exhibits a certain degree of robustness.

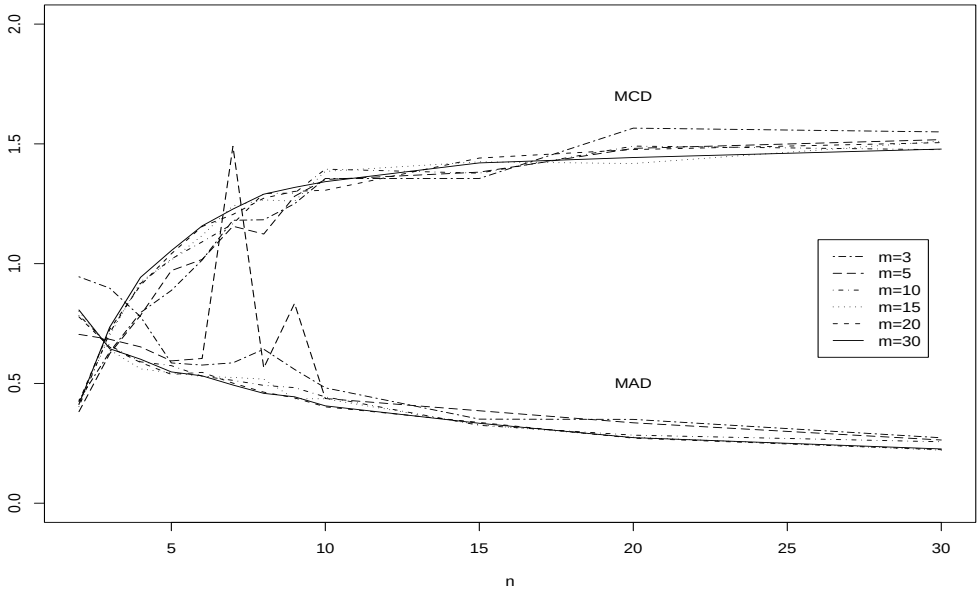
## Acknowledgements

This effort was jointly accomplished by the Boeing led team and the United States Government under the guidance of NAVAIR. This work was funded by DARPA/DSO and administered by NAVAIR under Technology Investment Agreement N00421-01-3-0098. The program acknowledges the guidance and support of Dr. Leo Christodoulou of DARPA/DSO for this effort. The technical monitor for the program is Dr. Ray Meilunas of NAVAIR. In addition, the authors acknowledge collaboration with Dr. Raj Rajagopal of The Boeing Company on this effort. The authors are indebted to Saveliy Uryasev at American Optimal Decisions, Inc., for expert programming assistance.

## References

Koenker, R., and Bassett, G. (1982). An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association*, *77*, 407–415.

Figure 1: Simulation results of the CVaR regression method applied to failure loads drawn from Weibull distributions. The plot shows Mean Absolute Deviation (MAD) between true and estimated 10th quantiles across batches ($m$), and Mean CVaR Deviation (MCD) for CVaR regression fits. The number of test points per batch ($n$) was also varied.





Koenker, R. (1994). Confidence intervals for quantile regression. *Proceedings of the 5th Prague Symposium on Asymptotic Statistics*, P. Mandl and M. Huskova (eds.). Heidelberg: Physica-Verlag.

Rockafellar, R., and Uryasev, S., (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*, 21–41.

Rockafellar, R.T. and Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking and Finance 26*(7), 1443–1471.

Rockafellar, R.T., Uryasev S., and Zabarankin, M. (2002). Deviation Measures in Generalized Linear Regression. *Research Report 2002-9*, ISE Dept., University of Florida.